International Journal of Computing and Digital Systems
ISSN (2210-142X)
Int. J. Com. Dig. Sys. 10, No.1 (Aug-2021)

http://dx.doi.org/10.12785/ijcds/100178

A Framework for Segmentation and Classification of Arrhythmia Using Novel Bidirectional LSTM Network

Archana Ratnaparkhi[1], Pallavi Deshpande[1] and Gauri Ghule[1]

[1] Dept. of Electronics and Telecommunication, Vishwakarma Institute of Information Technology, Maharashtra, India

Received 01 Jan. 2021, Revised 06 Apr. 2021, Accepted 15 Apr. 2021, Published 5 Aug. 2021

Abstract: Segmentation of ECG to obtain significant and relevant features has been a significant and inevitable step in achieving dimensionality reduction in automated heart disease diagnosis systems. Reduction in mortality due to the cardiovascular issues rate can only be achieved through accurate and rapid classification. Nonstationarity and high variability augment the complexity of the detection process in the domains of time and frequency. These difficulties are further enhanced due to imbalanced and vague datasets. In this paper, we propose to use a deep learning module to tackle the imbalance in the datasets by applying recurrent neural networks using long short-term memory layers (LSTM) to classify ECG into two classes. It has been seen that LSTM networks can effectively extract sequential timing information in the input ECG samples. To remove the imbalance in the datasets, oversampling and focal loss-based weight balancing techniques have been used, which eventually enhance the accuracy of classification. The proposed approach, an LSTM network with oversampling technique, provides an accuracy of 99.54%, which is considerably better than the traditional approaches that yield an accuracy of around 98%. Moreover, this method is insensitive to the quality of the ECG signal due to the fuzzification process followed in the initial stages of processing the dataset. Deployment of the proposed method for bio-signal telemetry or pharmaceutical research to assist physicians in their work is the most promising advancement in this domain.

Keywords: ECG, LSTM, Segmentation, Classification, Imbalanc

1. Introduction

An electrocardiogram (ECG) is the most significant indicator of heart conditions. According to a World Health Organization report, nearly 18% people die of cardiovascular diseases every year and a significant number of people lose their lives as a result of delayed medication. Automated diagnosis and analysis are indispensable for quick clinical diagnosis. Various methods of ECG classification include time domain methods which involve evaluation of intervals and amplitudes for feature set creation [1][2][3][4]. Additionally, frequency domain methods have been employed by [5][6],[7] to extract the significant frequencies that assist in heartbeat detection. A vast review of arrhythmia classification methods has been done by [8], [9]and [10]. Segmentation based on Markov models carried out by [11],[12][13] has suggested that these models are ill-suited for ECG feature extraction, and semi-Markov modelling needs to be carried out. The time-frequency approach has proven highly effective in extracting frequency at a finer resolution [14],[15][16]. SVM, radial basis function NN-based classifier [17], multilayer perceptron and various search algorithms have enhanced the performance of classifiers [18],[19]. Many of the methods discussed in the given literature employ hard-coded features for signal processing, segmentation and detection, which eventually leads to high false positives and consequently to misdiagnosis. We use a deep learning-based automated classification process to deal with the challenges involved in classification of ECG that involves high variability. Long term short memory (LSTM) is an advanced deep learning method [20] for processing time series. LSTM networks have been widely used for speech processing, natural language processing and handwriting recognition [21].

In this work, we propose using a bidirectional LSTM model for appropriate prediction of ECG samples. The novelty of this work lies in the use of a dual procedure to deal with the imbalance in the datasets. The efficacy of the proposed method is proven through extensive experimentation and comparison with existing feature extraction techniques. The next section provides a survey of the research carried out in using variants of the deep learning model.

E-mail: archana.ratnaparakhi@viit.ac.in, pallavi.deshpande@viit.ac.in, gauri.ghule@viit.ac.in

2. Related work

Few networks can be considered special neural networks where information is passed from the past event to its successor. Recurrent neural networks (RNN) are an example of such networks that help in analyzing short sequences. LSTM networks are redesigned RNN models which allow the relevant information to loop through a network and maintain a memory state. Significant information is retained while less significant information is eliminated form the process. Recently, LSTM-based deep learning on ECG was implemented by researchers to extract the temporal features. This model attained an accuracy of 99.86%. Yildirim et al. also carried out arrhythmia detection on long ECG segments with 91% accuracy. An auto encoder model based on LSTM networks has been developed by [22] for ECG detection. Effective error profile modelling using LSTM networks has been done by [23]. Evaluation of features for prediction of errors using LSTM in the form of overall summaries has been proposed and a successful predictor for the same was developed by [23] with good performance. [24] classified ECG beats with 74% accuracy. The proposed RNN-LSTM based ECG classification does not require any hand-crafted signals; it involves simplicity in the modelling for arrhythmia detection. The inherent presence of loops allows the network to store previous information about the temporal shape of ECG samples. Significant characteristic points can hence be retained and trivial ones can be lost. A primary problem with RNN is that long-term dependencies cannot be resolved. Vanishing gradient problem wherein weights having value less than 1 are multiplied several times so that they gradually reduce to zero without any significant change in the previous weights is a major issue with RNN structures. Sequences that show long-term dependency pose a problem for neural nets in their learning phase. LSTMs are specialized RNNs that are able to retain previous information for a longer duration. Fig. 1 shows the the internal structure of an LSTM cell.

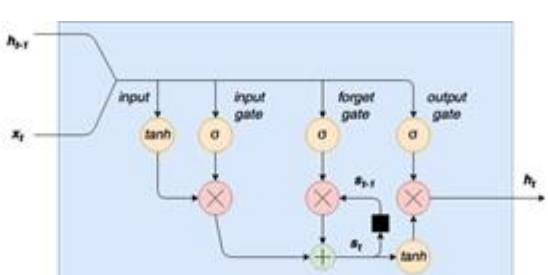

Figure 1. LSTM cell [25]

RNNs that look like interconnected chains as seen in Fig. 2 are the most appropriate choice while dealing with long sequences of ECG signals. The time-dependent nature of the network is modelled using LSTM networks. This is illustrated in Fig. 3. Theoretically, RNNs are certainly capable of retaining information termed as "long term dependencies". However, in practice, the performance degrades as dependencies become too long [27]. LSTM networks have been applied to a variety of problems and seem to work tremendously well [26]. The connected network seen in Fig. 4 is the LSTM structure that has repeating modules—just like RNNs with four neural network modules connected in a special way. Every line in Fig. 4 denotes a connection from input to output. The pink circles denote point-wise operations while yellow boxes represent a neural network.

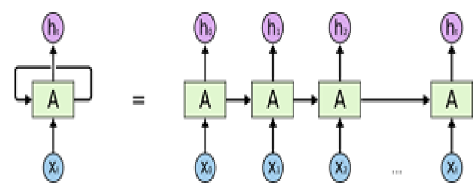

Figure 2. RNN cell [26]

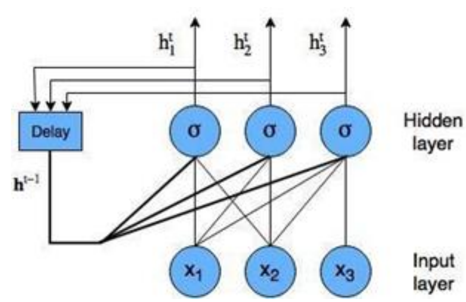

Figure 3. RNN cell [27]

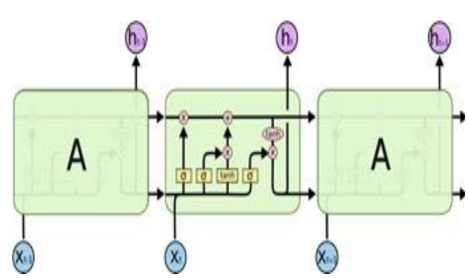

Figure 4. LSTM structure [26]

3. Rationale

Imbalance in the datasets adversely affects the classification accuracy. This happens primarily for two reasons: 1) as beats from one class occupy a large proportion of the total, the training efficiency goes down, and 2) Model degeneration takes place due to class imbalance. The problem of class imbalance has been addressed in literature using a cost-sensitive adaptation of the LSTM network. These cost-sensitive items have been introduced in back propagation learning procedures for classification purpose. The oversampling method has been employed to solve the problem of class imbalance. Specific probability distribution functions such as Gamma and Gaussian have been used by [28] to generate new data points. A few researchers have combined CNN and LSTM layers for diagnosis of multi-class arrhythmia to achieve an accuracy of 98%. Another method that deals with the class imbalance problem is the "focal loss method" which assigns lower weights to the attributes with higher contributions in terms of number and, more weight is assigned to samples that are fewer.This method has been used in our work as it assigns more weight to the samples which are hard to classify during the training process. Undersampling of samples belonging to majority class has also been a method worked upon by [29]. Synthetic minority oversampling with random undersampling has been employed by Liu et al. to deal with the class imbalance problem.

In the work proposed in this paper, we have used the focal loss method to classify the two class ECG samples. We have referred to the standard benchmark database which is widely used in literature—the MIT-BIH arrhythmia database—for experimentation. Novelty lies in use of a focal loss parameter to undermine the majority samples such that their contribution to the total loss is insignificant despite having these samples in large numbers. Basically, the algorithm trains the network with a sparse set of samples. The extensive experimentation on the database shows that results obtained outperform most of the previous results in terms of accuracy and complexity. Comprehensive experimentation with optimizers prominently highlights the efficacy of the LSTM network proposed in the paper.

4. Methodoogy

A detailed description of the modules involved in classification of ECG beats into Normal and Atrial Fibrillation beats is shown in Fig. 5. The input ECG samples are taken from Physionet Challenge Dataset [30]. Apart from 5050 Normal Beats and 738 Atrial Fibrillation beats, this database also contains noisy recordings beats. Hence, the preliminary denoising process has been carried out using enhanced savitzky golay filtering to have a cleaner ECG dataset. Every beat is 9000 samples long. A sequence length of 1000 samples is chosen as input to the bidirectional LSTM (BI-LSTM) network. Owing to its ability to acquire contextual information from past and future samples, the current sample is predicted

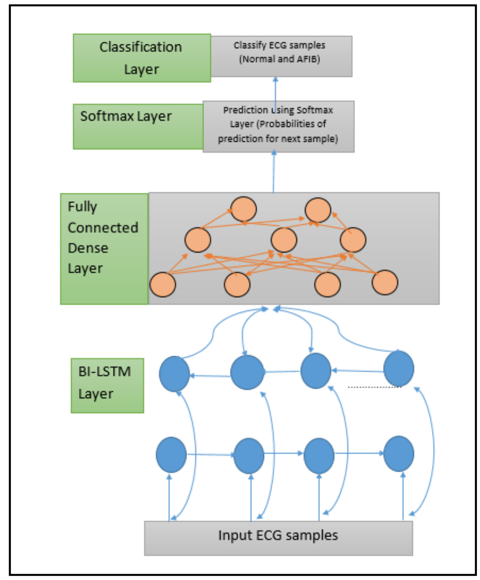

Figure 5. Methodology

This prediction is in the form of probabilities. As seen in Fig. 5, the BI-LSTM network is further connected to a fully connected dense structure of hidden neurons. The weights of these intermediate hidden neurons are altered using a backpropagation algorithm to have the softmax layer make a correct prediction. This layer maps all the inputs in the range of 0 and 1 which is similar to the generation of probabilities. A binary cross entropy based classification further classifies the samples into two. The flowchart given in Fig. 6 shows the various steps involved in the LSTM-based classification of ECG samples. Prediction of the beat from input samples using LSTM network has been carried out in the R environment using OSTSC package as well as in MATLAB. An oversampling method to classify uni-variant multi nominal time-series has been used for experimentation. Oversampling has been used marginally to equalize the number of samples in two classes. The focal loss method with variation in batch size, focusing parameter, and drop-out values has been extensively studied in this work to provide an analysis of the effect of variation of these parameters on classification accuracy. The novelty of this work lies in the use of LSTM networks with oversampling as well as the focal loss method for ECG classification. The results clearly indicate the efficacy of the above mentioned method which is discussed in the next section. The primary utility of the bidirectional LSTM model is that it can process the sequence in both directions whereas a regular LSTM network can only deal with the sequence in a forward direction.

Hence, in the proposed work, a bidirectional LSTM model has been used for classification of Normal and

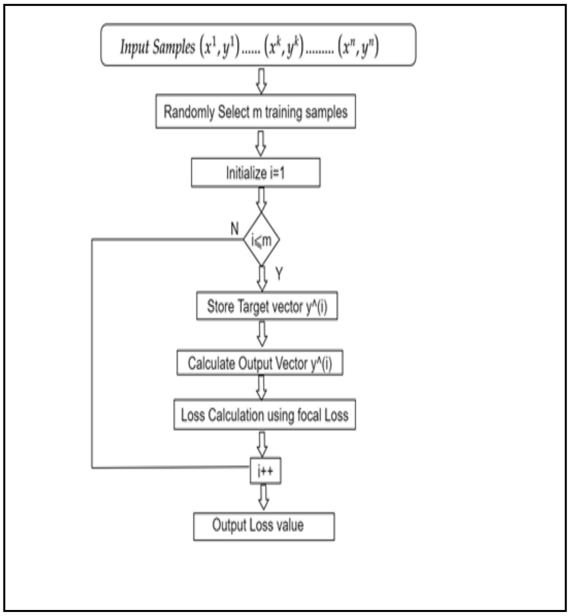

Figure 6. Flowchart

Atrial Fibrillation ECG signals using a novel method to deal with class imbalance.

5. EXPERIMENTATION

The dataset employed is the MIT BIH Arrhythmia dataset was obtained from the Physionet 2017 Challenge, which is available at [30].The sampling frequency of The ECG signals considered here is to be 300 Hz for classification of signals into Normal and Abnormal beats. The procedure involves binary classification of the input samples using the LSTM network in a deep learning framework. The evaluations have been done on an Intel core i7 processor in Microsoft Windows 10 64 bit operating system. Software simulations have been carried out in MATLAB and R environment. A histogram of the sample length as seen in Fig. 7 shows that most of the input signals are almost 9000 samples long. Visualization of the input samples as seen in Fig. 8 shows that atrial fibrillation samples are spaced irregularly while normal beats are seen regularly spaced. Generally, the P wave which occurs before QRS complex is absent in atrial fibrillation. A two-stage classification process is carried out to compare the effect of dataset imbalance on the classification accuracy using LSTM network. In the first stage, the raw ECG signals with 5050 Normal samples and 718 abnormal

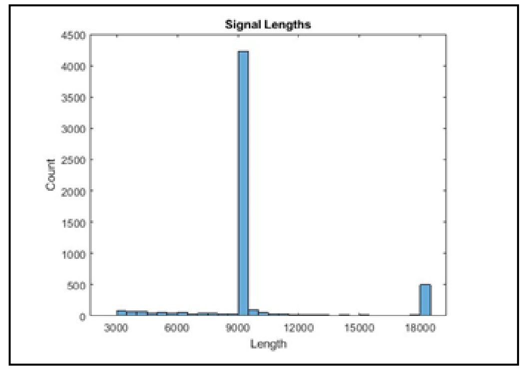

Figure 7. Histogram of Input samples

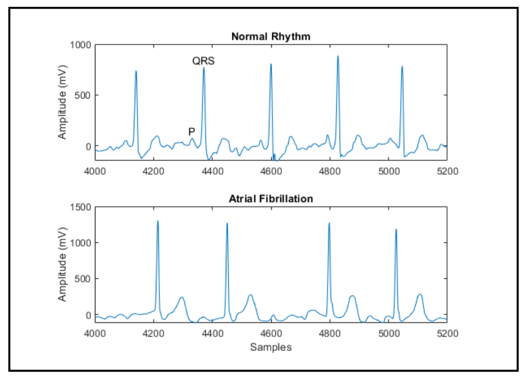

Figure 8. Input samples: Normal and Atrial Fibrillation samples

samples have been used for classification by splitting the dataset in the ratio 9:1 for train and test respectively. A 10-fold cross validation method has been used to compute the classification accuracy. The LSTM network is trained for the given dataset according to the configuration as indicated in Table 1.

A. Evaluation Metrics

The evaluation metrics based on the confusion matrix have been indicated by the formulae denoted by equations (1)–(5).

$$Accuracy(Ar) = \frac{Ps + Ns}{Ps + Ns + Fps + Fns} \quad (1)$$

$$Recall(Cr) = \frac{Ps}{Ps + Fns} \quad (2)$$

$$Precision(Pr) = \frac{Ps}{Ps + Fps} \quad (3)$$

$$Specificity(SP) = \frac{Ns}{Ns + Fps} \quad (4)$$

$$F1 = \frac{2 * Cr * Pr}{Cr + Pr} \quad (5)$$

Parameters such as Accuracy, Recall, Precision and F1 score parameters have been employed for an unbiased evaluation as the datasets are imbalanced in nature. Various cases, including change in batch size, change in dropout values and change in optimizers, have been studied to get an idea of the optimal LSTM configuration for ECG classification. Primarily, the effect of extracting features such as entropy and instantaneous frequency has also been studied and comparative results have been investigated.

B. Effect of Optimizers on Accuracy

The effect of various optimizers on the classification accuracy of the LSTM configuration has been studied and tabulated in Tables 2–3. Optimizers play a crucial role in updation of weights to minimize the loss cost. The momentum parameter, when used with optimizers, certainly accelerates the convergence process. Hence, the classification results with LSTM configuration based on stochastic gradient descent with momentum (SGDM) has been used for experimentation and compared with other optimizers. The training plot for the SGDM optimizer-based LSTM network as seen in Fig. 11 oscillates around 50% without trending either downward or upward. This indicates that we need to change the training option or optimizer.

Decreasing the batch size or the learning rate might help in convergence at the cost of longer training time. The Nesterov accelerated gradient algorithm is a slight modification to the classical stochastic gradient algorithm wherein the weight modification is done by looking at the future samples. Adagrad, another modified SGDM algorithm, has also been used for study. In this algorithm, frequently occurring parameters undergo fewer updates while infrequent parameters undergo more updates. Adagrad automates the selection of learning rate but the problem of radically diminishing learning rates exists, which is resolved in the Adam optimizer. The accuracy using Adam optimizer is 90% as seen in Fig. 12 and the focal loss curve over the epoch seems gradual. Hence, convergence is slow. Memory requirements of the Adam optimizer-based LSTM is also low. The additional momentum parameter when passed to the LSTM network configuration with the Adam optimizer speeds up the convergence as seen in Fig.14.The Adadelta optimizer, as seen in Fig.15, also tries to resolve the problem of aggressively reducing learning rate in Adagrad by restricting the window of the past accumulated gradient to a fixed size. The Nesterov-accelerated Adaptive Moment (Nadam) Estimation, as seen in Fig. 13, combines the classical stochastic gradient and Adam optimizer.

TABLE I. LSTM NETWORK CONFIGURATIONS

LSTM cells	*Network Layers*	*Optimizers*	*Drop Out*	*Epoch*	*Batch Size*	*Cost function*
64	4	Adam	0	350	150	Focal Loss
64	4	Adadelta	0	350	150	Focal Loss
64	4	Nadam	0	350	150	Focal Loss
64	4	RMSprop	0	350	150	Focal Loss
64	4	Adam	0	350	150	Oversampling with Focal Loss

TABLE II. CLASSIFICATION ACCURACY FOR DIFFERENT DROP OUT VALUES

Drop Out	*0*	*0.1*	*0.2*	*0.3*	*0.4*	*0.5*
Accuracy	98.24	97.52	99.21	98.76	98.54	99.54

TABLE III. CLASSIFICATION ACCURACY FOR DIFFERENT DROP OUT VALUES

Batch Sizes	*100*	*150*	*200*	*250*	*300*	*350*
Accuracy	98.24	97.52	99.21	98.76	98.54	99.54

Nadam is generally used for noisy and unstable gradients. Table 2 indicates the effect of change in dropout on the accuracy. Change in dropout value does not affect the classification accuracy considerably. It is seen highest at a dropout value of 0.5.

The effect of change in the focal loss parameter on the accuracy was evaluated. The effect of this parameter on abnormal beats was minor; however, as focal loss was increased further, the loss in correct detection of normal samples reduced. As focal loss was increased beyond 3, the misclassification showed a sharp increase. This can be seen in Table 3. Moreover, the experimentation has been done for various batch sizes as seen in Table 3. An optimal batch size of 150 yields the highest accuracy with due consideration given to the evaluation time. As the batch size increases, the execution time also increases exponentially. The effect of using features such as instantaneous frequency and entropy on classification accuracy can be seen in Fig. 9. These have been extracted using the spectogram on the MATLAB platform. The results indicate an accuracy of around 90% with a low computational complexity. Owing to the high variability in ECG signals and the instantaneous frequency, which is calculated from the short time frequency and transforms based on fixed window lengths, a generic high accuracy model cannot be devised. Hence, the proposed BI-LSTM network that caters to class imbalance using oversampling and focus loss generalizes well on the Physionet Challenge Dataset.

6. COMPARATIVE ANALYSIS

Classification of ECG samples has been a topic of research for decades. ECG being a highly non-stationary signal, extraction of significant information from the morphology becomes a difficult task. Support vector-based feature extraction and classification methods as used by Raj et al. and Sharma et al. have shown promising results. The proposed work based on BLSTM, however, certainly shows better accuracy and exhibits robustness due to novelty in removing the class imbalance. Oversampling to equalize the classes and focal loss method to correctly weigh the features has significantly increased the classification accuracy. The wavelets-based clustering method for ECG classification as shown in Table 5 exhibits an accuracy of around 96% at the cost of an increase in computational complexity. The work done by Oh et al. and Yildrim et al. as seen in Table 5 is comparable to the proposed work. As shown in Fig. 9, the confusion matrix was

TABLE IV. OVERALL ACCURACY FOR DIFFERENT VALUES OF FOCUSING PARAMETER (γ)

Focusing Parameter-(γ)	0	0.5	1	2	3	4
Accuracy	94.26	95.58	97.62	98.45	97.42	93.82

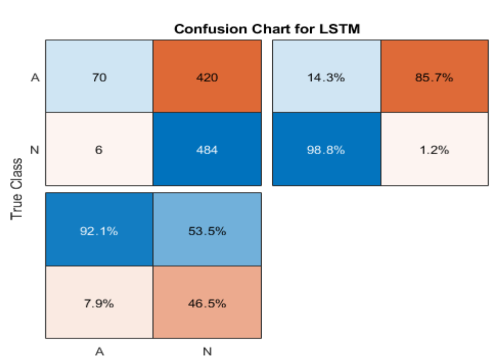

Figure 9. Accuracy and Focal Loss curves for proposed BILSTM model

Figure 10. Confusion Matrix for BILSTM with Feature Extraction module

TABLE V. COMPARATIVE ANALYSIS

Works	Year	Method	Accuracy	Recall	PP
Martis et al.	2013	PNN	99%	98.69%	99.9%
Raj et al.	2016	SVM-PSO	99.58%	-	-
Sharma et al.	2017	HHM-SVM	99.51%	89.64%	99.71%
Jung Lee et al.	2017	WKNN	96.12%	96.12%	99.91%
Yildrim et al.	2018	DU LSTM	99.25%	-	-
Oh et al.	2018	CNN LSTM	98.10%	97.50%	98.70%
Proposed Work	**2020**	**BI LSTM**	**99.54%**	**98.46%**	**99%**

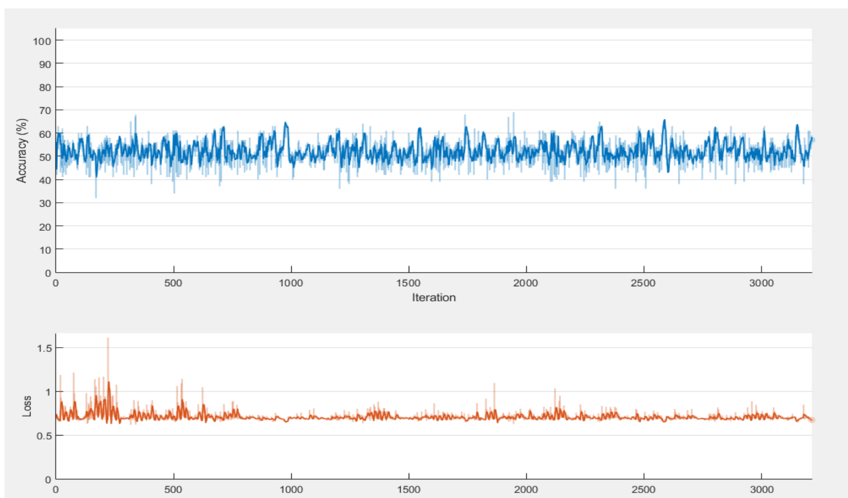

Figure 11. Accuracy and Focal Loss curves using SGDM optimizer

obtained for the input raw Physionet 2017 Challenge ECG dataset that contained rhythms other than normal and Atrial fibrillation. The performance metrics, sensitivity (92.1%), specificity (53.%), precision (14%) and F1 score (24.73%) clearly indicate low accuracy. To improve upon the classification accuracy, initial preprocessing in fuzzy rough (FR) domain [31] was carried out as datasets contains few noisy and unrelated samples. More explanation on preprocessing in the FR domain has been explained in [31]. Improved results have been shown in Fig. 10 which clearly indicates the efficacy of the proposed model as cross entropy loss almosts tends towards zero in lesser time compared to the previous model. The time required for training is considerably reduced to achieve nearly 99% classification accuracy.

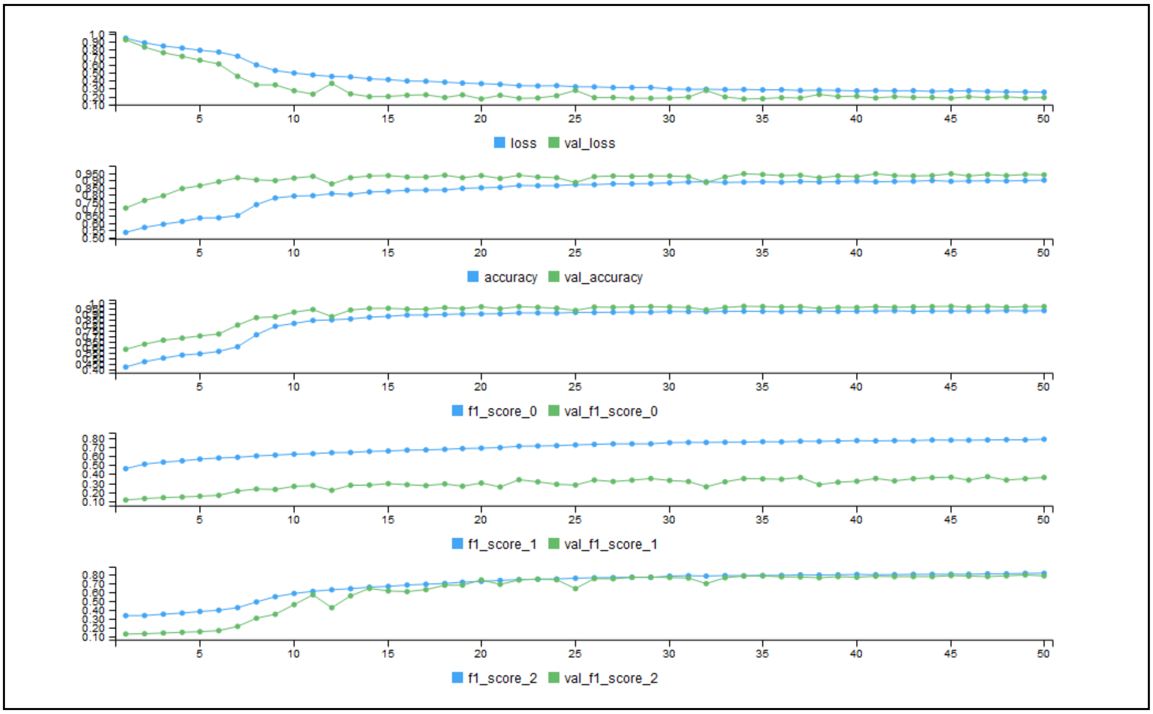

Figure 12. Accuracy and Focal Loss curves using Adam optimizer

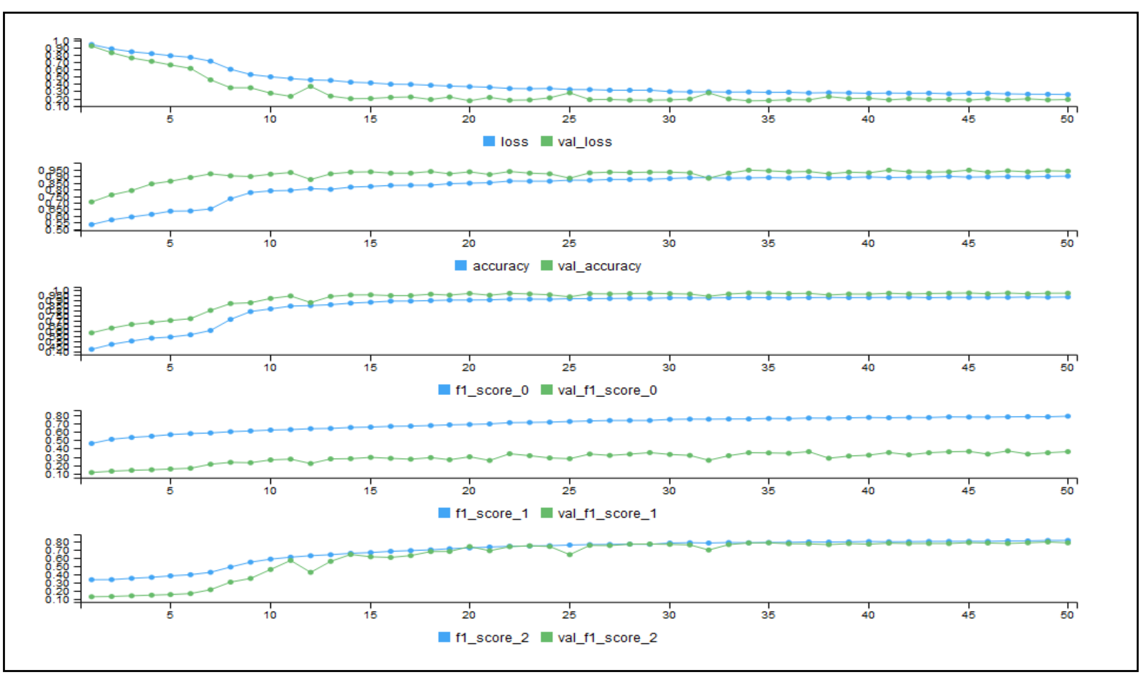

Figure 13. Accuracy and Focal Loss curves using Nadam optimizer

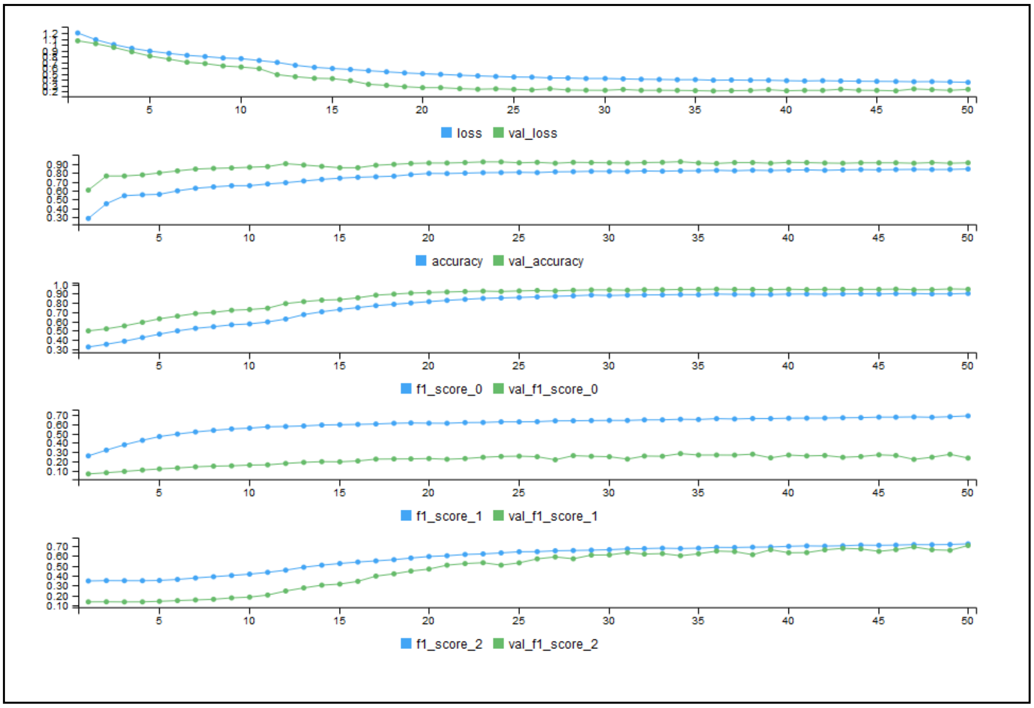

Figure 14. Accuracy and Focal Loss curves using Adam optimizer with momentum (RMRprop) optimizer

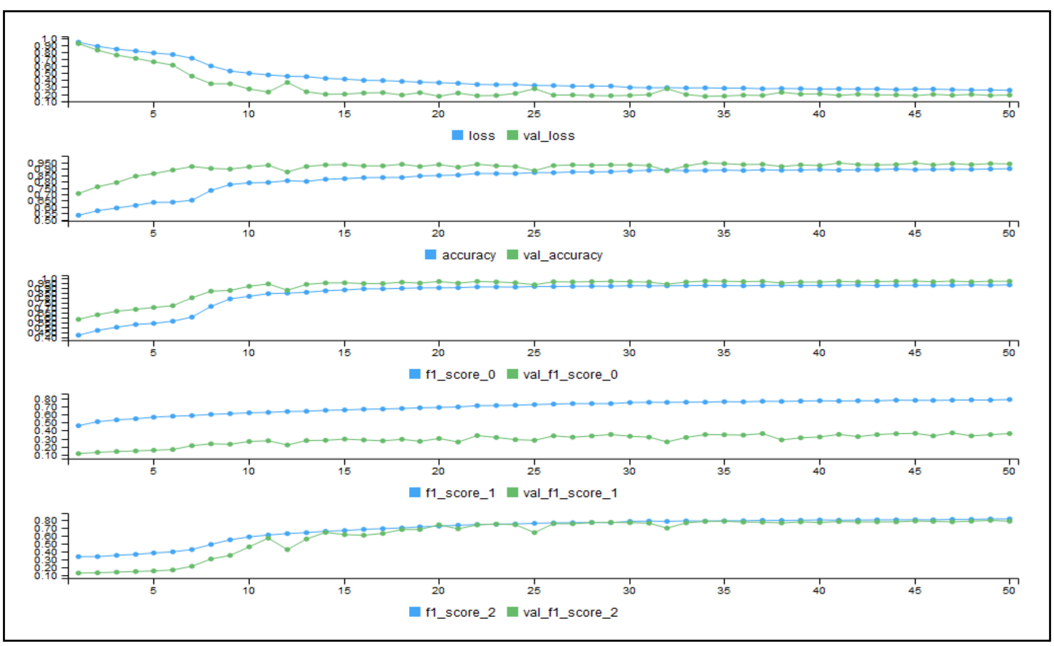

Figure 15. Accuracy and Focal Loss curves using Adadelta optimizer

7. Conclusion

In this study, we explore the use of LSTM networks for a binary classification of ECG signals. Various configurations of the LSTM network have been studied extensively to achieve better convergence. The experimentation has been carried out on MIT-BIH Arrhythmia datasets. The novelty of the work lies in effectively handling the imbalanced datasets to achieve an accuracy of 99.54% at a dropout value of 0.5 with the Adam optimizer. Variation in batch size and optimizers slightly affects the proposed network. The efficacy of the model is also tested by providing variation in focal loss values. This model can certainly be extended to include ECG beats other than arrhythmia beats. The robustness of the proposed model with variation in noise levels can be tested and would provide further research direction.

Acknowledgment

We are grateful to Savitribai Phule Pune University and Research wing, VIIT, for their technical and administrative support.

References

[1] P. De Chazel and R. Reilly, "A comparison of the ecg classification performance of different feature sets," in Computers in Cardiology 2000. Vol. 27 (Cat. 00CH37163). 1em plus 0.5em minus 0.4em IEEE, 2000, pp. 327–330.

[2] B. Anuradha, K. S. Kumar, and V. V. Reddy, "Classification of cardiac signals using time domain methods," ARPN Journal of Engineering and Applied Sciences, vol. 3, no. 3, pp. 7–12, 2008.

[3] D. Moses, "Real time cardiovascular disease diagnosis on mobile devices using novel methods for ecg compression extraction and classification," 2016.

[4] M. Kropf, D. Hayn, and G. Schreier, "Ecg classification based on time and frequency domain features using random forests," in 2017 Computing in Cardiology (CinC). 1em plus 0.5em minus 0.4em IEEE, 2017, pp. 1–4.

[5] C. H. Lin, "Frequency-domain features for ecg beat discrimination using grey relational analysis-based classifier," Computers & Mathematics with Applications, vol. 55, no. 4, pp. 680–690, 2008.

[6] R. Abeysekera and B. Boashash, "Time-frequency domain features of ecg signals: Their application in p wave detection using the cross wigner-ville distribution," in International Conference on Acoustics, Speech, and Signal Processing. 1em plus 0.5em minus 0.4em IEEE, 1989, pp. 1524–1527.

[7] A. Jafari, "Sleep apnoea detection from ecg using features extracted from reconstructed phase space and frequency domain," Biomedical Signal Processing and Control, vol. 8, no. 6, pp. 551–558, 2013.

[8] Z. Ebrahimi, M. Loni, M. Daneshtalab, and A. Gharehbaghi, "A review on deep learning methods for ecg arrhythmia classification," Expert Systems with Applications: X, p. 100033, 2020

[9] B. Pyakillya, N. Kazachenko, and N. Mikhailovsky, "Deep learning for ecg classification," in Journal of physics: conference series, vol. 913, no. 1. 1em plus 0.5em minus 0.4em IOP Publishing, 2017, p. 012004.

[10] A. Peimankar and S. Puthusserypady, "Dens-ecg: A deep learning approach for ecg signal delineation," Expert Systems with Applications, vol. 165, p. 113911, 2021.

[11] N. P. Hughes, L. Tarassenko, and S. J. Roberts, "Markov models for automated ecg interval analysis," in Advances in Neural Information Processing Systems, 2004, pp. 611–618.

[12] S. Graja and J. M. Boucher, "Hidden markov tree model applied to ecg delineation," IEEE Transactions on Instrumentation and Measurement, vol. 54, no. 6, pp. 2163–2168, 2005.

[13] A. Koski, "Modelling ecg signals with hidden markov models," Artificial intelligence in medicine, vol. 8, no. 5, pp. 453–471, 1996.

[14] J. P. Mart13 ̈053'fnez, R. Almeida, S. Olmos, A. P. Rocha, and P. Laguna, "A wavelet-based ecg delineator: evaluation on standard databases," IEEE Transactions on biomedical engineering, vol. 51, no. 4, pp. 570–581, 2004.

[15] S. Pal and M. Mitra, "Detection of ecg characteristic points using multiresolution wavelet analysis based selective coefficient method," Measurement, vol. 43, no. 2, pp. 255–261, 2010.

[16] R. D. Labati, E. Muñoz, V. Piuri, R. Sassi, and F. Scotti, "Deep-ecg: Convolutional neural networks for ecg biometric recognition," Pattern Recognition Letters, vol. 126, pp. 78–85, 2019.

[17] M. Akhbari, N. M. Ghahjaverestan, M. B. Shamsollahi, and C. Jutten, "Ecg fiducial point extraction using switching kalman filter," Computer methods and programs in biomedicine, vol. 157, pp. 129–136, 2018.

[18] P. Pławiak and U. R. Acharya, "Novel deep genetic ensemble of classifiers for arrhythmia detection using ecg signals," Neural Computing and Applications, pp. 1–25, 2019.

[19] J. F. Sedehi, N. J. Dabanloo, G. Attarodi, and M. E. Zadeh, "The prediction of ventricular fibrillation based upon hrv signal using combination of genetic algorithm and neural networks."

[20] U. B. Baloglu, M. Talo, O. Yildirim, R. San Tan, and U. R. Acharya, "Classification of myocardial infarction with multi-lead ecg signals and deep cnn," Pattern Recognition Letters, vol. 122, pp. 23–30, 2019.

[21] C. Wigington, S. Stewart, B. Davis, B. Barrett, B. Price, and S. Cohen, "Data augmentation for recognition of handwritten words and lines using a cnn-lstm network," in 2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR), vol. 1. 1em plus 0.5em minus 0.4em IEEE, 2017, pp. 639–645.

[22] B. Hou, J. Yang, P. Wang, and R. Yan, "Lstm based auto-encoder model for ecg arrhythmias classification," IEEE Transactions on Instrumentation and Measurement, 2019.

[23] S. Chauhan, L. Vig, and S. Ahmad, "Ecg anomaly class identification using lstm and error profile modeling," Computers in biology and medicine, vol. 109, pp. 14–21, 2019.

[24] A. Mostayed, J. Luo, X. Shu, and W. Wee, "Classification of 12-lead ecg signals with bi-directional lstm network," arXiv preprint arXiv:1811.02090, 2018.

[25] A. Thomas, Keras LSTM tutorial â€" How to easily build a powerful deep learning language model, Feb 3, 2020. [Online]. Available: https://adventuresinmachinelearning.com/keras-lstm-tutorial/ =0pt

[26] A. Ola, Understanding LSTM Networks, August 27, 2015. [Online]. Available: https://colah.github.io/posts/2015-08-Understanding-LSTMs/ =0pt

[27] S. Hochreiter, Y. Bengio, P. Frasconi, J. Schmidhuber et al., "Gradient flow in recurrent nets: the difficulty of learning long-term dependencies," 2001.

[28] H. R. Sanabila and W. Jatmiko, "Ensemble learning on large scale financial imbalanced data," in 2018 International Workshop on Big Data and Information Security(IWBIS). 1em plus 0.5em minus 0.4em IEEE, 2018, pp. 93–98.

[29] X. Wu, Z. Sui, C. H. Chu, and G. Huang, "Detection of atrial fibrillation from short ecg signals using a hybrid deep learning model," in International Conference on Smart Health. 1em plus 0.5em minus 0.4em Springer, 2019, pp. 269–282.

[30] G. D. Clifford, C. Liu, B. Moody, H. L. Li-wei, I. Silva, Q. Li, A. Johnson, and R. G. Mark, "Af classification from a short single lead ecg recording: the physionet/computing in cardiology challenge 2017," in 2017 Computing in Cardiology (CinC). 1em plus 0.5em minus 0.4em IEEE, 2017, pp. 1–4.

[31] A Ratnaparkhi, D. Bormane, and R. Ghongade, "Performance analysis of fuzzy rough assisted classification and segmentation of paper ecg using mutual information and dependency metric," in International Conference on ISMAC in Computational Vision and Bio-Engineering. 1em plus 0.5em minus 0.4em Springer, 2018, pp. 789–801.

Archana Ratnaparkhi has completed her Bachelor of Engineering in Electronics from K.I.T.S Ramtek in 2001 with First Class. She has done her Masters in Electronics Engineering from Shivaji University in 2005. Currently, she is working as Assistant Professor in Vishwakarma Institute of Information Technology. She was awarded a Ph.D. in Biomedical Engineering from Savitribai Phule Pune University in 2020. She has done extensive research in ECG Signal Processing and Analysis. Her research interest also includes the study of Rough sets and Fuzzy rough domain. She has 3 patents and 1 copyright to her credit and has been working closely with hospitals through her projects with social health care in rural areas. Her association with professional bodies includes a membership with the IEEE Women Affinity Group and Young Professional Group

Pallavi Deshpande is working in the field of Biometric Hand Recognition for the last 8 years. She has done her graduation and post-graduation from Savitribai Phule Pune University. Her doctoral work has added numerous awards and accolades to her record

Gauri Ghule has completed her graduation and post-graduation from Savitribai Phule Pune University. She has patents and copyrights in the domain of Speech Recognition and Processing. She has published numerous papers in revered journals in this regard.